

Classification of scale free networks

K.-I. Goh*, E.S. Oh*, H. Jeong†, B. Kahng*, and D. Kim*

*School of Physics and Center for Theoretical Physics, Seoul National University, Seoul 151-747, Korea

†Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea

While the emergence of a power law degree distribution in complex networks is intriguing, the degree exponent is not universal. Here we show that the betweenness centrality displays a power-law distribution with an exponent η which is robust and use it to classify the scale-free networks. We have observed two universality classes with $\eta \approx 2.2(1)$ and 2.0 , respectively. Real world networks for the former are the protein interaction networks, the metabolic networks for eukaryotes and bacteria, and the co-authorship network, and those for the latter one are the Internet, the world-wide web, and the metabolic networks for archaea. Distinct features of the mass-distance relation, generic topology of geodesics and resilience under attack of the two classes are identified. Various model networks also belong to either of the two classes while their degree exponents are tunable.

Emergence of a power law in the degree distribution $P_D(k) \sim k^{-\gamma}$ in complex networks is an interesting self-organized phenomenon in complex systems (1; 2; 3). Here, the degree k means the number of edges incident upon a given vertex. Such a network is called scale-free (SF) (4). Real world networks which are SF include the author collaboration network (5) in social systems, the protein interaction network (PIN) (6) and the metabolic network (7) in biological systems, and the Internet (8) and the world-wide web (WWW) (9; 10) in communication systems. The power-law behavior means that most vertices are sparsely connected, while a few vertices are intensively connected to many others and play an important role in functionality. While the emergence of such a SF behavior in degree distribution itself is surprising, the degree exponent γ is not universal and depends on the detail of network structure. As listed in Table 1, numerical values of the exponent γ for various systems are diverse but most of them are in the range $2 < \gamma \leq 3$. From the viewpoint of theoretical physics, it would be interesting to search a universal quantity associated with SF networks.

Recently a physical quantity called “load” was introduced as a candidate for the universal quantity in SF networks. It quantifies the load of a vertex in the transport of data packet along the shortest pathways in SF networks (11). It was shown that the load distribution exhibits a power law, $P_L(\ell) \sim \ell^{-\delta}$, and the exponent δ is robust as $\delta \approx 2.2$ for diverse SF networks with various degree exponents in the range $2 < \gamma \leq 3$. Since the universal behavior of the load exponent was obtained empirically, fundamental questions such as how the load exponent is robust in association with network topology or the possibility of any other universal classes existing, have not been explored yet. In this paper, we address those issues in detail.

While the load is a dynamic quantity, it is closely related to a static quantity, the “betweenness centrality (BC)”, commonly used in sociology to quantify how much a given person is influential in a society (12). To

be specific, BC is defined as follows. Let us consider the set of the shortest pathways, or geodesics, between a pair of vertices (i, j) and denote their number by $C(i, j)$. Among them, the number of the shortest pathways running through a vertex k is denoted by $C_k(i, j)$. The fraction $g_k(i, j) = C_k(i, j)/C(i, j)$ may be interpreted as the amount of the role played by the vertex k in social relation between two persons i and j . Then the BC of the vertex k is defined as the accumulated sum of $g_k(i, j)$ over all ordered pairs for which a geodesic exists, *i.e.*,

$$g_k = \sum_{i \neq j} g_k(i, j) = \sum_{i \neq j} \frac{C_k(i, j)}{C(i, j)}. \quad (1)$$

Because of only slight difference between load and BC, both quantities behave very closely. In fact, the BC g_k of each vertex is exactly the same as the load for tree graphs. In general, distributions of the two are indistinguishable within available resolutions. The BC distribution follows a power law,

$$P_B(g) \sim g^{-\eta}, \quad (2)$$

where g means BC and the exponent η is the same as the load exponent δ . Since the topological feature of a network can be grasped more transparently using BC, we deal with BC in this work.

Based on numerical measurements of the BC exponent for a variety of SF networks, we find that SF networks can be classified into only two classes, say, class I and II. For the class I, the BC exponent is $\eta \approx 2.2(1)$ and for the class II, it is $\eta \approx 2.0(1)$. We conjecture the BC exponent for the class II to be exactly $\eta = 2$ since it can be derived analytically for simple models. We show that such different universal behaviors in the BC distribution originate from different generic topological features of networks. Moreover, we study a physical problem, the resilience of networks under an attack, showing different behaviors for each class, as a result of such difference in generic topologies. It is found that the networks of class II are much more vulnerable to the attack than those for class I.

To obtain our results, we use available data for real world networks or existing algorithms for model networks. Once a SF network is constructed, we select a pair of vertices (i, j) on the network and identify the shortest pathways between them. Next, BC is measured on each vertex along the shortest pathways using the modified version of the breath-first search algorithm introduced by Newman (13). We measure the BC distribution and the exponent η for a variety of networks both in real world and in silico.

Real World and Artificial Networks Investigated

The networks that we find to belong to the class I with $\eta = 2.2(1)$ include: (i) The co-authorship network in the field of the neuroscience, published in the period 1991-1998 (14), where vertices represent scientists and they are connected if they wrote a paper together. (ii) The protein interaction network of the yeast *S. cerevisiae* compiled by Jeong *et al.* (6) (PIN1), where vertices represent proteins and the two proteins are connected if they interact¹. (iii) The core of protein interaction network of the yeast *S. cerevisiae* obtained by Ito *et al.*² (PIN2) (15). (iv) The metabolic networks for 5 species of eukaryotes and 32 species of bacteria in Ref.(7), where vertices represent substrates and they are connected if a reaction occurs between two substrates via enzymes. The reaction normally occurs in one direction, so that the network is directed. (v) The Barabási-Albert (BA) model (16) when the number of incident edges of an incoming vertex $m \geq 2$. (vi) The geometric growth model by Huberman and Adamic (10). (vii) The copying model (17), whose degree exponent is in the range of $2 < \gamma \leq 3$. (viii) The undirected or the directed static model (11), whose degree exponent is in the range of $2 < \gamma \leq 3$ or $2 < (\gamma_{\text{in}}, \gamma_{\text{out}}) \leq 3$, respectively. (ix) The accelerated growth model proposed by Dorogovtsev *et al.* (18). (x) The fitness model (19) with a flat fitness distribution. (xi) The stochastic model for the protein interaction networks introduced by Solé *et al.* (20). All those networks (i)-(xi) exhibit a power-law behavior in the BC distribution with the exponent $\eta \approx 2.2(1)$. Detailed properties of each network are listed in Table 1. The representative BC distributions for real world networks (i), (iii), and (iv) are shown in Fig.1a.

The networks that we find to belong to the class II with $\eta = 2.0$ include: (xii) The Internet at the autonomous systems (AS) level as of October, 2001 (22). (xiii) The metabolic networks for 6 species of archaea

in Ref.(7). (xiv) The WWW of nd.edu (9). (xv) The BA model with $m = 1$ (16). (xvi) The deterministic model by Jung *et al.* (23). In particular, the networks (xv) and (xvi) are of tree structure, where the edge BC distribution can be solved analytically. The detailed properties of each network are listed in Table 1. The BC distributions for real world networks (xii) and (xiv) are shown in Fig.1b. Since the BC exponents of each class are very close numerically, one may wonder if there exist really two different universal classes apart from error bar. To make this point clear, we plot the BC distributions for the BA model with $m = 1, 2$ and 3 in Fig.2, obtained from large system size, $N = 3 \times 10^5$. We can see clearly different behaviors between the two BC distributions for the cases of $m = 1$ (class II) and of $m = 2$ and 3 (class I).

Topology of the Shortest Pathways

To understand the generic topological features of the networks in each class, we particularly focus on the topology of the shortest pathways between two vertices separated by a distance d . Along the shortest pathways, we count the total number of vertices $\mathcal{M}(d)$ lying on these roads, averaged over all pairs of vertices separated by the same distance d . Adopting from the fractal theory, $\mathcal{M}(d)$ is called the “mass-distance” relation. We find that it behaves in different ways for each class; For the class I, $\mathcal{M}(d)$ behaves nonlinearly (Figs.3a-b), while for the class II, it is roughly linear (Fig.3c-d).

For the networks belonging to the class I such as the PIN2 (iii) and the metabolic network for eukaryotes (iv), $\mathcal{M}(d)$ exhibits a non-monotonic behavior (Fig.3a-b), *viz.*, it exhibits a hump at $d_h \approx 10$ for (iii) or $d_h \approx 14$ for (iv). To understand why such a hump arises, we visualize the topology of the shortest pathways between a pair of vertices, taken from the metabolic network of a eukaryote organism, *Emericella nidulans* (EN), as a prototypical example for the class I. Fig.4a shows such a graph with linear size 26 edges ($d = 26$), where an edge between a substrate and an enzyme is taken as the unit of length. From Fig.4a, one can see that there exists a blob structure inside which vertices are multiply connected, while vertices outside are singly connected. What is characteristic for the class I is that the blob is localized in a small region. To see this, we measure the mass density $m(r; d)$, the average number of substrates or enzymes located at position r ($\sum_{r=1}^d m(r; d) = \mathcal{M}(d)$). The average is taken over all possible pairs of vertices (56 pairs), separated by the same distance $d = 26$. Note that the metabolic network is directed, so that the position r is uniquely defined. As shown in Fig.5, we find that $m(r; d)$ is sharply peaked at $r = 3$, and is larger than 1 only at $r = 2, 4$, and 6 for substrates. Thus the blob structure is present, even after taking averages, and is localized in a small region of size $d_b \simeq 4 \sim 5$, centered at almost the same position $r \approx 3 \sim 4$ for different pairs of vertices. The blob size d_b can be measured in another way. In a given graph of the shortest pathways, we delete singly connected substrates successively until none is left and measure the linear size

¹ The network is composed of disconnected clusters of different sizes, *viz.*, small isolated clusters as well as a giant cluster. For both (ii) and (xi), the degree distribution is likely to follow a power law but there needs an exponential cutoff to describe its tail behavior for finite system. However, it converges to a clean power law for (xi) as system size increases, but the converging rate is rather slow (21). Despite this abnormal behavior in the degree distribution for finite system, the BC distribution follows a pure power law with the exponent $\eta \approx 2.2(1)$ in (ii) and (xi).

² In contrast to (ii), the degree distribution obeys a power law.

of the remaining structure. When averaged over all pairs of vertices with separations $d > 10$, it comes out to be $d_b \approx 4.5$, well consistent with the value obtained previously for $d = 26$ only. Due to this blob structure, the mass-distance relation increases abruptly across $d = 4$ as shown in Fig.3b.

Next, we measure the average mass of blob, that is, the number of vertices inside a blob for a given graph of the shortest pathways with separation d , averaged over all pairs of vertices with the same separation. We find that the average blob mass is broadly distributed in the range $3 < m_b < 23$. In particular, relatively heavy blob masses, $m_b = 15 \sim 23$, mainly comes from the graphs whose linear size is $d = 8 \sim 14$. Due to those blobs with heavy mass, the mass-distance relation exhibits a hump, and decreases at around $d = 14 \sim 16$, beyond which, the mass $\mathcal{M}(d)$ increases linearly by the presence of singly connected vertices. In short, the anomalous behavior in the mass-distance relation is due to the presence of a *compact and localized* blob structure in the topology of the shortest pathways between a pair of vertices for the metabolic network of eukaryotes. We have checked the mass-distance relations and the graphs of the shortest pathways for other networks belonging to the class I such as the PIN2 and metabolic networks for other organisms, and found that such topological features are generic, generating the anomalous behavior in the mass-distance relation. It still remains a challenge to derive the BC exponent $\eta \approx 2.2$ analytically from such structures.

For the class II, the mass depends on distance linearly, $\mathcal{M}(d) \sim Ad$ for large d (Fig.3c-d). Despite the linear dependence, the shortest pathway topology for the case of $A > 1$ is more complicated than that of the simple tree structure where $A = 1$. Therefore, the SF networks in the class II are subdivided into two types, called the class IIa and IIb, respectively. For the class IIa, $A > 1$ and the topology of the shortest pathways includes multiply connected vertices (Figs.4b and 4c), while for the class IIb, $A \sim 1$ and the shortest pathway is almost singly connected (Fig.4d). Examples in real world networks in the class IIa are the Internet at the AS level ($A \sim 4.5$) and the metabolic network for archaea ($A \sim 2.0$), while that in the class IIb is the WWW ($A \sim 1.0$).

Let us examine the topological features of the shortest pathways for the networks in the class IIa and IIb more closely. First, for the class IIa, we visualize in Fig.4b a shortest pathway in the Internet system between a pair of vertices separated by 10 edges, the farthest separation. It contains a blob structure, but the blob is rather extended as $d_b = 5$, comparable to the maximum separation $d = 10$. We obtain $d_b = 5$ for $d = 11$ for another system. For comparison, $d_b \approx 4.5$ for $d = 26$ in the class I. Moreover, the feature-less mass-position dependence $m(r; d)$ we have found implies that while most blobs are located almost in the middle of the shortest pathways, which seems to be caused by the geometric effect, there are a finite number of blobs located at the verge of the shortest pathways. Note that $m(r; d) = m(d - r; d)$ since the

Internet is undirected. Owing to the extended structure and the scattered location of the blob, the mass-distance relation exhibits the linear behavior, $\mathcal{M}(d) \sim Ad$ with $A \approx 4.5$. The extended blob structure is also observed in the metabolic network for archaea (Fig.4c). Since the network in this case is directed, the symmetry in $m(r; d)$ does not hold. However, the blob structure extends to almost one half of maximum separation, and the shortest pathways are very diverse, so that their topological property such as the mass-distance relation $\mathcal{M}(d)$ is similar to that of the Internet.

The WWW is an example belonging to the class IIb. For this network, the mass-distance relation exhibits $\mathcal{M}(d) \sim 1.0d$, suggesting that the topology of the shortest pathway is almost singly connected, which is confirmed in Fig.4d. When a SF network is of tree structure, one can solve the distribution of BC running through each edge analytically, and obtain the BC exponent to be $\eta = 2$. A derivation of this exact result is presented in the Appendix.

Comparison of the Resilience under Attack

So far we have investigated the topological features of the shortest pathways of SF networks of each class. Then what would be distinct physical phenomena originated from such different topological features? Associated with this question, we investigate a problem of the resilience of network under a malicious attack. It is known that SF networks are extremely vulnerable to the intentional attack to a few vertices with high degree, while it is very robust to random failures (24; 25). To compare how vulnerable a network in each different class is under such attacks, we first construct a directed network whose numbers of vertices and edges, and the degree distribution are identical to those of the WWW (xiv), but whose BC exponent is 2.2. It can be generated, for example, by following the stochastic rule introduced in the directed static model (11). For both the WWW in real world and the artificial model network, we remove vertices in the descending order of BC successively. As vertices are removed, both the mean distance $\langle d \rangle$ between two vertices, known as the diameter, and the relative size of the giant cluster S are measured as a function of the fraction of removed vertices f . As can be seen in Fig.6a, the diameter of the WWW with $\eta = 2.0$ (class IIb) increases more rapidly than that with $\eta = 2.2$ (class I) and shows discrete jumps while vertices are removed. Also the relative size of the largest cluster decreases more rapidly for $\eta = 2.0$ than for $\eta \approx 2.2$ (Fig.6b). This behavior arises from the fact that the shortest pathway consists of mainly singly connected vertices for the class IIb, so that there is no alternative pathways with the same distance when a single vertex lying on the shortest pathway is removed. For the Internet in real world with $\eta \approx 2.0$ in the class IIa and an artificial network with $\eta \approx 2.2$ with the same numbers of vertices and edges and the identical degree distribution, the differences in the diameter $\langle d \rangle$ and in

the relative size S of the largest cluster appear to be rather small (Figs.6c-d), in comparison to the case of the WWW (Figs.6a-b). This is because the shortest pathways are multiply connected for the class IIa.

Conclusions

In conclusion, we have found that the betweenness centrality can determine the universal behavior of SF networks. By examining a variety of real world and artificial SF networks, we have observed two distinct universality classes whose BC exponents are $\eta \simeq 2.2(1)$ (class I) and 2.0 (class II), respectively. The mass-distance relation is introduced to characterize the topological features of the shortest pathways. It shows a hump for the class I networks due to compact and localized blobs in the shortest pathway topologies, while it is roughly linear for the class II ones which are more or less tree-like. The class II networks can further be divided into two types depending on whether the shortest pathway topology contains diversified pathways (class IIa) or mostly singly connected ones (class IIb). Distinct features of the resilience under attack arising from the different topologies of the shortest pathways are also identified. Since SF networks show the small world property, the topology of the shortest pathways should be of relevance for characterizing the network geometry. Indeed the mass-distance relations for different universality classes show different behaviors. Such a relation between the universality class and the topological features of the shortest pathways may be understood from the perspective of the fact that the geometric fractal structure of the magnetic domains in equilibrium spin systems at criticality can classify the universality classes. Further characterizations in static and dynamic properties and possible evolutionary origin of the universality classes are interesting questions left for future study.

Acknowledgments

This work was supported by the Korean Research Foundation (Grant No. 01-041-D00061) and BK21 program of MOE, Korea.

APPENDIX

Here we present the analytic derivation of the BC distribution for a tree structure, however, the derivation is carried out for the edge BC rather than the vertex version³. The edge betweenness centrality is defined on edges as in Eq.(1), with the subscript k now denoting a bond. Without any rigorous proof, we assume that the distributions of vertex BC and edge BC behave in the same manner particularly on tree structures, which is confirmed by numerical simulations. We have also checked the identity

between the vertex BC and the edge BC distributions for a deterministic model of scale-free tree introduced by Jung *et al.* (23), which will be published elsewhere.

We consider a growing tree network such as the BA type model with $m = 1$, where a newly introduced vertex attaches an edge to an already existing vertex j with the probability proportional to its degree as $(k_j + a) / \sum_{\ell} (k_{\ell} + a)$. Then the network consists of $N(t) = t + 1$ vertices and $L(t) = t$ edges at time t . The stationary degree distribution is of a power law with $\gamma = 3 + a$ (26; 27). Each edge of a tree divides the vertices into two groups attached to either sides of the edge. Let $P_s(m, t)$ be the probability that the edge born at time s bridges a cluster with m vertices on the descendant side and another with remaining $t + 1 - m$ vertices on the ancestor side. Due to the tree structure, the BC running through that edge born at s is given as $g = 2m(t + 1 - m)$, independent of the birth time s . The probability $P_s(m, t)$ evolves as a new vertex attaches to one of the two clusters. The rate equation for this process is written as

$$P_s(m, t + 1) = r_1(m, t)P_s(m, t) + r_2(m - 1, t)P_s(m - 1, t), \quad (\text{A1})$$

where $r_1(m, t)$ is the probability that a new vertex attaches to the cluster with $(t + 1 - m)$ vertices on the ancestor side, and $r_2(m - 1, t)$ with $(m - 1)$ vertices on the descendant side. They are given explicitly as

$$r_1(m, t) = 1 - r_2(m, t) = \frac{(2 + a)(t - m) + at + 1}{2t + a(t + 1)}. \quad (\text{A2})$$

Since the amount of the BC on the edge s is independent of the birth time, we introduce $P(m, t)$,

$$P(m, t) = \frac{1}{t} \sum_{s=1}^t P_s(m, t), \quad (\text{A3})$$

which is the probability for a certain edge to locate between two clusters with m and $t + 1 - m$ vertices averaged over its birth time. The BC on that edge is still given by $2m(t + 1 - m)$. In terms of $P(m, t)$, Eq.(A1) can be written as

$$(t + 1)P(m, t + 1) = r_1(m, t)tP(m, t) + r_2(m - 1, t)tP(m - 1, t) \quad (\text{A4})$$

In the limit of $t \rightarrow \infty$, one may rewrite $P(m, t)$ in a scaling form, $P(m, t) = \mathcal{P}(m/t)$ and then Eq.(A4) is rewritten as

$$(t + 1)\mathcal{P}(x) - t\mathcal{P}(x) \simeq -x \frac{d\mathcal{P}(x)}{dx} - \mathcal{P}(x) \quad (\text{A5})$$

where $x = m/t$ and the approximation $\mathcal{P}(x - 1/t) \simeq \mathcal{P}(x) - (1/t)d\mathcal{P}(x)/dx$ has been used. From this we obtain that

$$\mathcal{P}(x) \sim \frac{1}{x^2}, \quad (\text{A6})$$

³ See Szabó *et al.* (cond-mat/0203278, 2nd version) for a mean field treatment of the vertex BC problem on trees.

independent of the tuning parameter a . Using $g = 2(t + 1 - m)m \sim 2t^2x$ for large t and finite m , Eq.(A6) becomes

$$P_B(g) \sim \frac{1}{g^2}. \quad (\text{A7})$$

Thus $\eta = 2$ is obtained for the tree structure, independent of $\gamma > 2$. General finite size scaling relations for $P_B(g)$ are discussed in Ref.(28).

References

- [1] Strogatz, S.H. (2001) *Nature (London)* **410**, 268-276.
- [2] Albert, R. & Barabási, A.-L. (2002) *Rev. Mod. Phys.* **74**, 47-97.
- [3] Dorogovtsev, S.N. & Mendes, J.F.F. (2002) *Adv. Phys.* **51**, 1079-1187.
- [4] Barabási, A.-L., Albert, R. & Jeong, H. (1999) *Physica A* **272**, 173-187.
- [5] Newman, M.E.J. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 404-409.
- [6] Jeong, H., Mason, S.P., Barabási, A.-L. & Oltvai, Z.N. (2001) *Nature (London)* **411**, 41-42.
- [7] Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N. & Barabási, A.-L. (2000) *Nature (London)* **407**, 651-654.
- [8] Faloutsos, M., Faloutsos, P. & Faloutsos, C. (1999) *Comp. Comm. Rev.* **29**, 251-262.
- [9] Albert, R., Jeong, H., & Barabási, A.-L. (1999) *Nature (London)* **401**, 130-131.
- [10] Huberman, B.A. & Adamic, L.A. (1999) *Nature (London)* **401**, 131-131.
- [11] Goh, K.-I., Kahng, B. & Kim, D. (2001) *Phys. Rev. Lett.* **87**, 278701.
- [12] Freeman, L.C. (1977) *Sociometry* **40**, 35-41.
- [13] Newman, M.E.J. (2001) *Phys. Rev. E* **64**, 016131; *ibid.* **64**, 016132.
- [14] Barabási, A.-L., Jeong, H., Ravasz, R., Neda, Z., Vicsek, T. & Schubert, A. (2002) *Physica A* **311**, 590-614.
- [15] Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. & Sakaki, Y. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 4569-4574.
- [16] Barabási, A.-L. & Albert, R. (1999) *Science* **286**, 509-512.
- [17] Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tomkins, A. & Upfal, E. (2000) in *Proc. IEEE FOCS 2000* (IEEE Computer Society Press, Los Alamitos, CA), pp. 57-65.
- [18] Dorogovtsev, S.N. & Mendes, J.F.F. (2001) *Phys. Rev. E* **63**, 025101(R).
- [19] Bianconi, G. & Barabási, A.-L. (2001) *Europhys. Lett.* **54**, 436-442.
- [20] Solé, R., Pastor-Satorras, R., Smith, E. & Kepler, T. (2002) *Advances in Complex Systems* **5**, 43-54.
- [21] Kim, J., Krapivsky, P.L., Kahng, B. & Redner, S. (2002) (cond-mat/0203167).
- [22] Meyer, D. (2001) *University of Oregon Route Views Archive Project* (<http://archive.routeviews.org>).
- [23] Jung, S., Kim, S. & Kahng, B. (2002) *Phys. Rev. E* **65**, 056101.
- [24] Albert, R., Jeong, H. & Barabási, A.-L. (2000) *Nature (London)* **406**, 378-381.
- [25] Cohen, R., Erez, K., ben-Avraham, D. & Havlin, S. (2000) *Phys. Rev. Lett.* **85**, 4626-4629; (2001) *ibid.* **86**, 3682-3685.
- [26] Krapivsky, P.L., Redner, S. & Leyvraz, F. (2000) *Phys. Rev. Lett.* **85**, 4629-4632.
- [27] Dorogovtsev, S.N., Mendes, J.F.F. & Samukhin, A.N. (2000) *Phys. Rev. Lett.* **85**, 4633-4636.
- [28] Goh, K.-I., Kahng, B. & Kim, D. (2002) *Physica A* (in press).

Figure Legends

Fig. 1 The BC distributions of real world networks.

(a) Networks belonging to the class I: Co-authorship network (\times), Core of PIN of yeast by Ito *et al.* (+), and metabolic network of *EN* (\diamond). The solid line is a fitted line with a slope -2.2 . (b) Networks belonging to the class II: WWW of *nd.edu* (\circ) and Internet AS as of October, 2001 (\square). The solid line has a slope -2.0 .

Fig. 2 Comparison of the BC distributions for the two classes.

BA model with $m = 1, 2$, and 3 are simulated for large system size, $N = 3 \times 10^5$, averaged over 10 configurations. The dotted line has a slope -2.0 and the dashed one, -2.2 .

Fig. 3 The mass-distance relation $\mathcal{M}(d)$.

(a) Core of PIN of yeast obtained by Ito *et al.*. (b) Metabolic networks of eukaryotes. Data are averaged over all 5 organisms in Ref.(7). Note that in this case we count only substrates for $\mathcal{M}(d)$. (c) Internet AS as of October, 2001. (d) WWW of *nd.edu*.

Fig. 4 Topology of the shortest pathways.

(a) The metabolic network of *EN* (eukaryote) of length 26. (b) The Internet AS of length 10. (c) The metabolic network of *Methanococcus jannaschii* (archae) of length 20. (d) WWW of *nd.edu* of length 20. In (a) and (c), circles denote substrates and rectangles denote intermediate states.

Fig. 5 The mass density.

$m(r; d)$ for *EN* with $d = 26$. Circles denote substrates and rectangles intermediate states.

Fig. 6 Attack vulnerability of the scale-free networks.

The WWW ($\eta = 2.0$) (\blacksquare) and the artificial directed SF network with $\eta = 2.2$ (\square), the Internet ($\eta = 2.0$) (\bullet) and the artificial undirected SF network with $\eta = 2.2$ (\circ): Changes in network diameter (a, c) and the relative size

of the largest cluster (b, d) are shown as a function of f , the fraction of removed vertices measured in percent (%).

Table Legend

Table 1 Natures of diverse SF networks.

Tabulated for each network are the size N , the mean degree $\langle k \rangle$, the degree exponent γ , and the betweenness centrality exponent η .

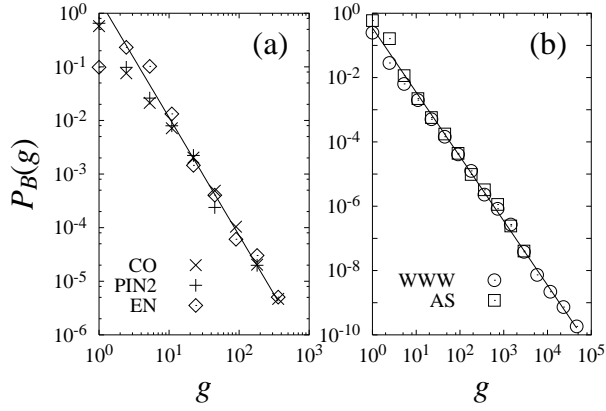


FIG. 1 Goh et al.

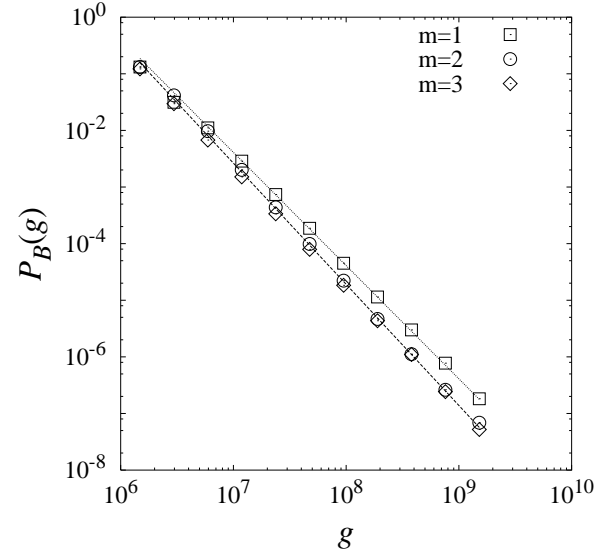


FIG. 2 Goh et al.

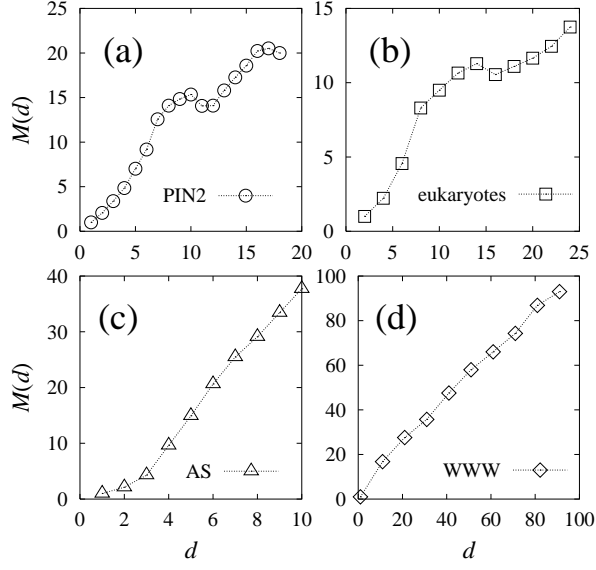


FIG. 3 Goh et al.

TABLE I Goh et al.

Class	System	N	$\langle k \rangle$	γ	η	Ref.
Class I	(i) Co-authorship	205 202	11.8	2.2(1)	2.2(1)	(14)
	(ii) PIN1	1846	2.39	2.4 (exp. cut-off)	2.2(2)	(6)
	(iii) PIN2	797	1.96	2.7(1)	2.2(1)	(15)
	(iv) Metabolic (eukaryotes, bacteria)	$\sim 10^3$	2-4	2.0-2.4	2.2(1)	(7)
	(v) BA model ($m \geq 2$)	3×10^5	$2m$	2.0-3.0	2.2(1)	(16)
	(vi) HA model	10^5	$\mathcal{O}(1)$	3.0(1)	2.2(1)	(10)
	(vii) Copying model	10^4	4	2.0-3.0	2.2(1)	(17)
	(viii) Static model	10^4	4,6,8	2.0-3.0	2.2(1)	(11)
	(ix) Accelerated growth model	10^4	$\mathcal{O}(1)$	3.0(1)	2.2(1)	(18)
	(x) Fitness model	10^4	4	2.25	2.2(1)	(19)
	(xi) PIN model	10^4	~ 2	2-3	2.2(1)	(20)
Class IIa	(xii) Internet AS	12 058	4.16	2.2	2.0(1)	(22)
	(xiii) Metabolic (archaea)	$\sim 10^3$	2-4	2.0-2.3	2.0(1)	(7)
Class IIb	(xiv) WWW	325 729	4.51	2.1/2.45	2.0	(9)
	(xv) BA tree ($m = 1$)	∞	2	> 2.0	2.0	(16)
	(xvi) Deterministic tree	∞	2	> 2.0	2.0	(23)

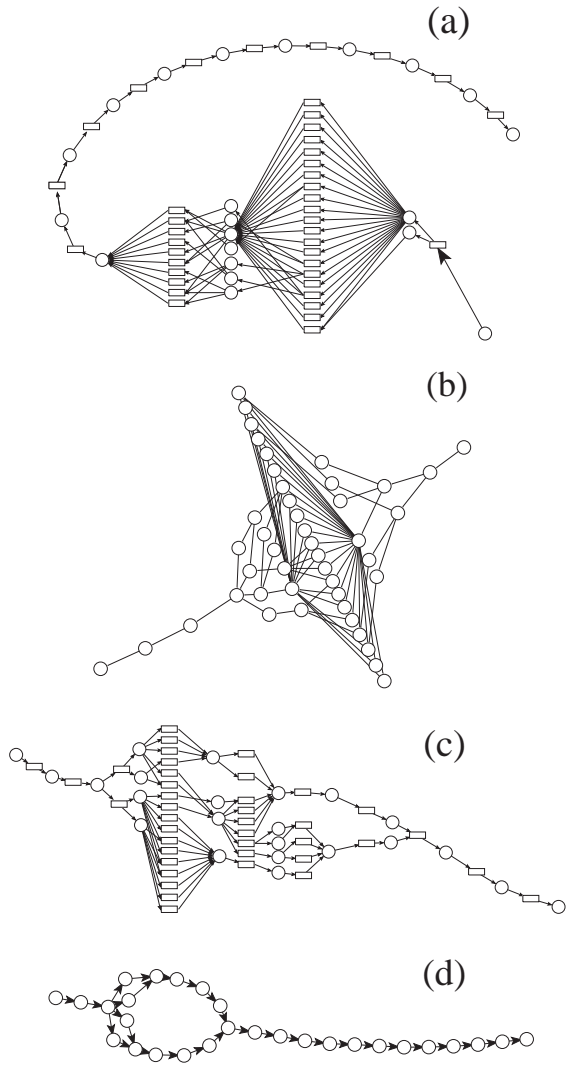


FIG. 4 Goh et al.

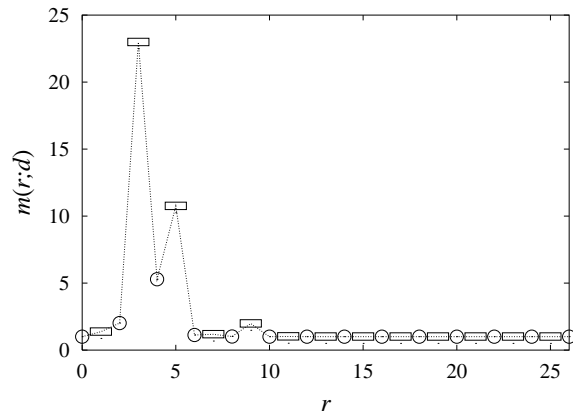


FIG. 5 Goh et al.

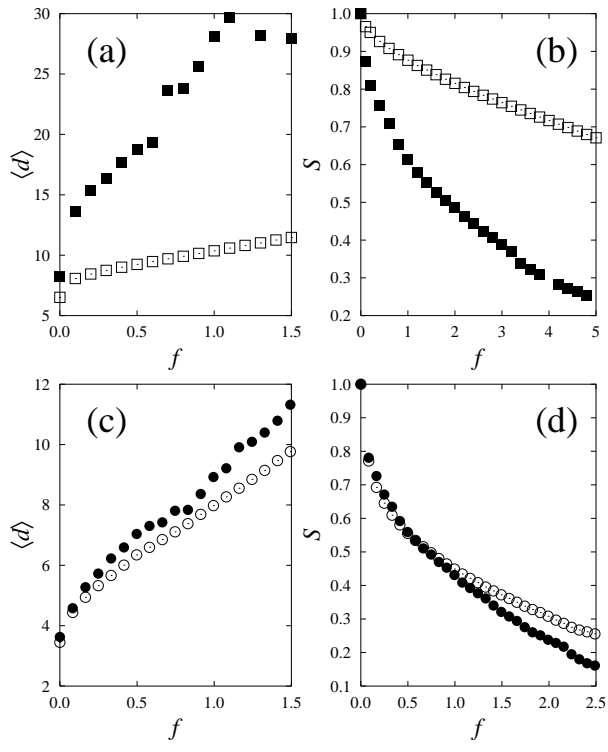


FIG. 6 Goh et al.